PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :
C12N 15/52, 15/53, 9/00
C12N 9/04, C12P 7/42

A1

(11) International Publication Number: WO 93/15208

(43) International Publication Date: 5 August 1993 (05.08.93)

(21) International Application Number: PCT/GB93/00204

(22) International Filing Date: 29 January 1993 (29.01.93)

(30) Priority data:
9202033.8 30 January 1992 (30.01.92) GB
9204702.6 4 March 1992 (04.03.92) GB

(71) Applicant *(for all designated States except US)*: GENZYME LIMITED [GB/GB]; 37 Hollands Road, Haverhill, Suffolk CB9 8PU (GB).

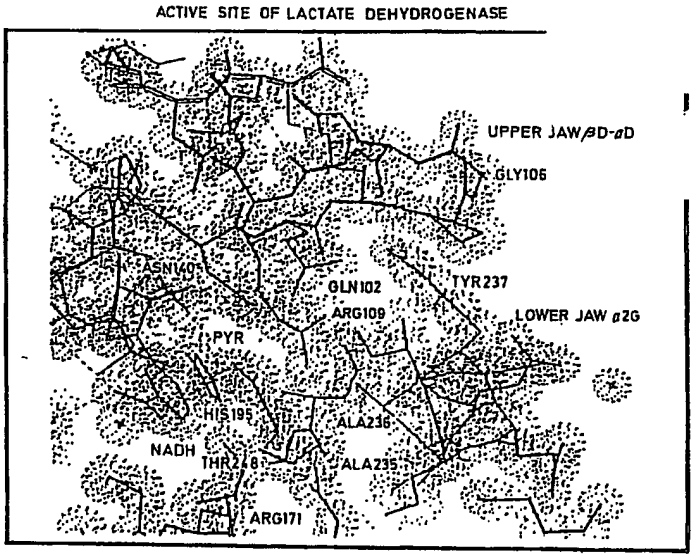
(72) Inventors; and
(75) Inventors/Applicants *(for US only)* : WILKS, Helen, Margaret [US/US]; 705E Mountain Road, Charlottesville, VA 22903 (US). HOLBROOK, Joseph, John [GB/GB]; Sleight Farm, Timsbury, Bath BA1 (GB). HART, Keith, William [GB/GB]; 36 Darnley Avenue, Orfield, Bristol BS7 0BS (GB). ELHAWRANI, Ayman [GB/GB]; Department of Biochemistry, University of Bristol, Bristol BS8 1TD (GB).

(74) Agents: FROUD, Clive et al.; Elkington and Fife, Prospect House, 8 Pembroke Road, Sevenoaks, Kent TN13 1XR (GB).

(81) Designated States: AT, AU, BB, BG, BR, CA, CH, DE, DK, ES, FI, GB, HU, JP, KP, KR, LK, LU, MG, MN, MW, NL, NO, PL, RO, RU, SD, SE, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: CHIRAL SYNTHESIS WITH MODIFIED ENZYMES



(57) Abstract

A method for modifying the specificity and/or efficiency of an enzyme, while retaining its catalytic activity, characterised in that it comprises: selecting an enzyme, the tertiary structure of which is substantially known or deduced; identifying at least one specificity and/or efficiency-related region; identifying or constructing unique restriction sites bounding the identified region in the DNA coding therefor; generating a DNA sequence which corresponds to at least a portion of the identified region, except that the nucleotides of at least one codon are randomized, or selecting as a substitute for at least a portion of the identified region an alternative such region, which may itself be similarly randomized; using the generated or substitute DNA sequence to replace the original such sequence; expressing the DNA including the generated or substitute DNA sequence; and selecting for a desired modification so that the DNA coding therefor may be isolated is disclosed.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FR	France	MR	Mauritania
AU	Australia	GA	Gabon	MW	Malawi
BB	Barbados	GB	United Kingdom	NL	Netherlands
BE	Belgium	GN	Guinea	NO	Norway
BF	Burkina Faso	GR	Greece	NZ	New Zealand
BG	Bulgaria	HU	Hungary	PL	Poland
BJ	Benin	IE	Ireland	PT	Portugal
BR	Brazil	IT	Italy	RO	Romania
CA	Canada	JP	Japan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo			SE	Sweden
CH	Switzerland	KR	Republic of Korea	SK	Slovak Republic
CI	Côte d'Ivoire	KZ	Kazakhstan	SN	Senegal
CM	Cameroon	LI	Liechtenstein	SU	Soviet Union
CS	Czechoslovakia	LK	Sri Lanka	TD	Chad
CZ	Czech Republic	LU	Luxembourg	TG	Togo
DE	Germany	MC	Monaco	UA	Ukraine
DK	Denmark	MG	Madagascar	US	United States of America
ES	Spain	ML	Mali	VN	Viet Nam
FI	Finland	MN	Mongolia		

CHIRAL SYNTHESIS WITH MODIFIED ENZYMES.

This invention relates to chiral synthesis; more particularly, it relates to the modification of enzymes to facilitate such synthesis.

Enzymes are biological catalysts which are specific both in terms of chemical activity and substrate structure, and it is this specificity which has been exploited in a variety of commercial applications. Although many such activities are known, it may be desirable to change the range of substrates that are suitable for catalysis and/or to change the efficiency of a given catalysis for a particular type of enzyme. Given a type of enzyme with known key elements vis-à-vis substrate preference and hence activity, it may be possible purposefully to change those elements to bring about desired modifications and hence to expand the potential industrial utility of a particular enzyme.

Enzyme activity is primarily controlled by the amino acid composition especially in certain important functional areas of the enzyme, altering these amino acids is known to change activity and may be achieved by the use of either specific or non-specific techniques. For example, the introduction of a neutralising amino acid may facilitate the catalysis of a substrate with an altered charge and this could be regarded as a predictable alteration, although no result may ever be predicted with total certainty, especially where the tertiary structures of enzymes are not as precisely known as would be necessary for complete confidence. However, while it is possible to make individual changes by known means, this would prove an almost infinite task and so it is often convenient initially to make a "macro-change" and then to "fine tune" with discrete changes. Of course, in a given case, a macro-change may prove to be sufficient, or, indeed, discrete changes may be all that are required.

Although alteration of the enzyme structure has been described, this is not achieved by any direct effect on the amino acid components, but by known techniques on the DNA encoding for the enzyme prior to protein transcription. Taking as an example the enzyme lactate dehydrogenase (natural substrate pyruvate), when acting on the carboxylic acid analogue of pyruvate, oxalo acetic acid, it would have substantially reduced activity due to the negative charge introduced into the active site. In this case, site-directed mutagenesis involving the introduction of a neutralizing charge into the correct region of the active site alters substrate specificity allowing the enzyme to take on the activity that would be expected of a malate dehydrogenase. Such specific mutations may be considered predictable in gross terms, but are very unlikely to be the ultimate refinement in increasing specificity towards such a substrate. For alternative substrates, such as those with increased alkyl chain lengths, phenyl residues or heterocyclic additions, predictions of site-specific changes are unlikely to be reliable. It is probable that the changes necessary to accommodate such "unnatural substrates" are most likely to be required adjacent to or in the active site region of the enzyme, which in many enzymes may involve up to 20 amino acids, which may be derived from many disparate parts of the primary sequence. Clearly, if one tried to proceed by alterations in individual amino acids, the scale of the undertaking would be impractical even with modern techniques.

In order to achieve the desired objective while circumventing the above disadvantages, it is possible in the case of lactate dehydrogenase, for example, to make use of the known loop region forming part of the active site. As a convenient first step, at least a portion of the loop region may be exchanged for a larger or smaller section of loop region from a similar enzyme. This may be expected to allow some variation in substrate specificity and relative catalytic efficiency, while retaining the typical activity. Having chosen the most promising loop region for a desired substrate, which could indeed be the

starting wild-type loop, specific amino acid residues may be targeted for further change. In order to secure the best possible option, it is necessary to survey all possible amino acid combinations in the positions of interest. This is done by generating random nucleotides in the region coding for the amino acids targeted. Following routine cloning, it becomes necessary to select for a desired modification from amongst the numerous alternatives produced. Such screens are in common use. This approach to enzyme engineering is facilitated by the introduction of unique endonuclease restriction sites into the coding DNA, if such are not already present, at desired points. Such changes may often be achieved by alteration in the bases without altering the amino acid encoded due to code degeneracy or alternatively they are achieved by the introduction of codes as far as possible for similar amino acids. This allows the region of particular interest to be handled independently of the remainder.

As will be appreciated from the foregoing, the present invention relates to a method for modifying the specificity and/or efficiency of an enzyme, while retaining its catalytic activity, characterised in that it comprises: selecting an enzyme, the tertiary structure of which is substantially known or deduced; identifying at least one specificity and/or efficiency-related region; identifying or constructing unique restriction sites bounding the identified region in the DNA coding therefor; generating a DNA sequence which corresponds to at least a portion of the identified region, except that the nucleotides of at least one codon are randomized, or selecting as a substitute for at least a portion of the identified region an alternative such region, which may itself be similarly randomized; using the generated or substitute DNA sequence to replace the original such sequence; expressing the DNA including the generated or substitute DNA sequence; and selecting for a desired modification so that the DNA coding therefor may be isolated.

It will be described in more detail below, but the

present method may be illustrated by reference to a dehydrogenase, in particular an α-hydroxy acid dehydrogenase, such as lactate dehydrogenase. In this illustration, it is the loop region of the enzyme which is identified initially as being specificity and/or efficiency-related. Generally, the randomized DNA is generated by means of an inosine triphosphate PCR method or a spiked oligonucleotide method or a PCR assembly method, all of which will be discussed in more detail below. If a substitute is to be selected for at least a portion of the region of interest, it is often based on a corresponding sequence from a similar enzyme. Once the original DNA sequence has been replaced by the generated or substitute DNA sequence, it is cloned into a plasmid or phage vector and transformed into a bacterium or virus for expression. Thereafter, a screen may be used to select for a desired modification. Taking L-lactate dehydrogenase as an example, positions 101 and 102 are particularly appropriate for randomization.

The present invention also relates to the use of such modified enzymes particularly in the production of chiral products. Often, such processes involve the use of a cofactor recycling system. One example is the reduction of 2-oxo-4-phenyl-propanoic acid characterised in that it comprises the use of L-lactate dehydrogenase which has been modified in the loop region by the present method and another is the reduction of 4-methyl-2-oxo-3-pentenoic acid characterised in that it comprises the use of MVS/GG obtainable by the present method.

Having outlined the present invention, it will now be described more fully.

The use of enzymes in chemical synthesis has gained increasing acceptance as an academic possibility, while its introduction into industrial chemical procedures is rare. The potential advantages of enzymes as catalysts, such as obtaining stereospecificity and regiospecificity under mild conditions, have initiated many attempts to obtain enzymes suitable for

particular chemical conversions.

Several approaches to selection of the enzyme are possible. Experimentation with currently-available enzymes may yield surprising results in terms of breadth of substrate specificity not predictable from the literature. It is thus possible to utilise commercially-available enzymes, which may have a low catalytic efficiency, but, because of cost, may form the basis of an industrial process. A second approach is to screen large numbers of environmental micro-organisms in an attempt to effect a particular transformation. Should such an activity be obtained, it is often required that the enzyme be obtained in a purer form than whole microbial cells or crude preparations thereof. To obtain enzymes from such a screen in sufficient quantity and at a reasonable cost for an industrial process requires extensive development often with the involvement of cloning and over-expression of the gene. Another approach for obtaining suitable enzyme catalysts is to modify the structure of an existing enzyme to improve its catalysis for a particular substrate. This approach of so-called "enzyme engineering", which is in its very early stages has great potential for the preparation of catalysts for the synthesis of homochiral molecules. The importance of these molecules in the synthesis of single isomer pharmaceuticals and agrochemicals is well recognised.

Despite the obvious attraction of enzyme engineering, the results of amino acid changes are often, at best, only of limited predictability due to the structural complexity of enzymes. At present, it is not possible to predict the effect of certain amino acid changes on the finer points of substrate recognition and catalytic performance where the substrate is altered in size and additional functionalities introduced from the natural substrate. It is generally easy to predict the removal of activity by the elimination of one of the catalytically-vital amino acids which are generally well known from the classical studies of enzyme mechanism and function. To

enhance the activity towards an unnatural substrate remains a challenge.

The opportunity for enzyme engineering may be calculated for a 300 residue protein of 20 amino acids as 10^{390} possible sequences. The vast majority of these sequences cannot have been explored for biological function. It may be suggested that a typical large protein of 300 amino acids residues cannot represent a global optimum for any biological function, but at best is an assembly of empirically optimised 25-35 amino acid domains. Thus, enzyme engineering should be capable of improving a large frame-work for any particular target function.

Recently, methods have been developed to express random sequences of DNA as protein fused to phage M13 coat protein and it has been suggested that it will be possible to mimic the process of evolution by suitable affinity chromatography to isolate both the required protein sequences and its gene (Kang, PNAS, 88, 1991, 4363). However, just as evolution has been unable to sample all possible sequences, so too the protein engineer will be limited to the number of M13 phage that may be screened (10^{15} plaque-forming units are produced per litre culture of E. coli cells containing the phage M13). With 10^{15} variants, the length of DNA which may be optimised is obtained from $4^N = 10^{15}$, i.e. N = 24 bases or 8 amino acids. The other problem encountered is that a phage display system determines binding not catalysis and thus is not designed to obtain enzymes with new chemical potential.

Random mutagenesis of existing proteins is also limited in its ability to produce radically altered proteins by problems of sampling all the possible variants. In addition, the genetic code is very resistant to change. Not only are codons redundant at the third position, but also amino acid residues with similar properties are coded by similar sequences and thus resistant to sparse mutagenesis. For example: (i) a codon having a T at the second position always codes for an amino acid residue having a

hydrophobic side chain; (ii) the codons for aspartate and glutamate differ only at the third position. Therefore, strategies, such as use of thioate nucleotides (Holm, Prot Eng, 3, 1990, 181), which create randomly dispersed mutations (in which only one mutation is likely to be present in any codon) are unlikely to yield new proteins having dramatically different properties to those of the parent proteins.

Although it should be possible to engineer any designed property into any protein framework, only those which have been well characterised are likely to be redesigned successfully.

In order to obtain the fundamental knowledge required for rational redesign, a combination of crystallography, site-directed mutagenesis and transient kinetic techniques was used to relate function to structure in the NAD-dependent lactate dehydrogenases from both prokaryotes and eukaryotes. That knowledge not only revealed those amino acids required for the catalytic pathway, but also mapped those amino acids which are part of a major rearrangement of shape which is induced when the negatively-charged substrate acid enters the active site and causes the protein to sequester the substrate in an internal vacuole which is sensitive to the size of the substrate and which contains exactly balanced charge. Using this knowledge, it has been possible to design specific new enzymatic properties with respect to charged substrates and so avoid the low statistical probabilities associated with random mutagenesis. It should, of course, be appreciated that the present invention is more generally applicable than to this particular illustration.

Accordingly Fig. 1 depicts the active site of lactate dehydrogenase. In this illustration, some of the residues which determine substrate specificity are carried on the under-surface of the "upper jaw". The rate-limiting step in lactate dehydrogenase catalysis is the rate at which this loop may sweep through a viscous solvent to close onto the upper surface of helix α2G. The rate-limiting step is largely independent of the

sequence of amino acids on the "upper jaw" and since the chemical step is much faster than the shape change, the lactate dehydrogenase system has the advantage that the loop sequence may be easily varied to achieve different substrate specificities without much danger that the chemical step will become rate-limiting. Thus, in order to obtain enzymes improved by engineering towards particular substrates, a combination of techniques may be preferentially employed. Specific residues may be changed to accommodate functional groups, such as an altered charge to that of the natural substrate, but to perfect the enzyme for activity towards a different substrate, elements of the infinite variability of random amino acid changes may be required. This may be applied to a particular area of the enzyme and selected for using screening techniques.

An object of the present invention was to modify an already useful, but substrate-restricted enzyme, S lactate dehydrogenase, to provide an improved catalyst for reduction of the α-keto group in acids larger than the natural substrate, pyruvate. In particular, the substrates of interest contain bulky aromatic groups.

The natural enzyme used as the basis for engineering was the thermophilic lactate dehydrogenase (LDH) isolated from Bacillus stearothermophilus, which has been cloned and expressed in Escherichia coli.

This enzyme has been one of the most thoroughly characterised protein frameworks (Dunn, C. R., et al, Philos. Trans. R. Soc. London Ser. B, 1991, 332, 184), including the study of inhibition, substrate interaction and genetic manipulation. The physical stability of the enzyme, especially to thermal denaturation, makes it an ideal candidate for demonstrating the features of redesign which would be generally applicable to α-hydroxy acid dehydrogenases, for example.

The modification of wild-type enzymes presents a significant challenge because, even in the case of a protein with considerable literature knowledge, the results may be unexpected and surprising. Thus, redesign of even well- studied enzymes is of limited predictability.

Changes in the amino acid composition of enzymes and thus effects on kinetics and substrate specificity have occurred throughout nature and various methods have been developed in order to potentiate the natural divergence of enzyme structure. Random mutations may be produced in genetic information (and thus in the protein coded for) by the use of classical mutagenesis. Lately, the technique of site directed mutagenesis has allowed the alteration of specific bases in genes, thus producing directed amino acid changes in the target protein at a known position. Using similar techniques, it has been possible to achieve the replacement of significant amino acid sequences in a functionally important area of the enzyme.

Detailed knowledge of the protein, such as primary sequence and tertiary structure from X-ray analysis, along with molecular modelling allow the identification of the position of various amino acids in what are known as conserved regions. This is illustrated with the nomenclature of the amino acids of various lactate dehydrogenase enzymes. Thus, any structure in the protein which is retained between species is regarded as conserved and probably essential for the enzyme's function. This information will allow any change in a particular enzyme to be pinpointed for all other homologous enzymes across all general substrate types; if this were not possible the enzymes would not fulfil the same biochemical function. The enzymes of particular interest at present are α-hydroxy acid dehydrogenases, which catalyse the NADH/NADPH dependent reduction of a keto group in an α-position to a carboxylic acid, or, alternatively, the reverse reaction where the α-hydroxy group is oxidised to the ketone.

Attempts to modify the enzyme lactate dehydrogenase to expand the natural substrate specificity to allow an increased reaction rate with larger substrates with various functional groups has led to the present unpredictable observations. Although it may be possible to prepare substrates and corresponding chiral products of interest by chemical synthesis, followed by wild-type enzyme reduction, such an approach may not be attractive and it may be that preparation via a redesigned protein framework may provide a more rational and cost effective approach. Additionally, the alteration of the enzyme has demonstrated that the activity towards the natural substrate may be so dramatically reduced that completely different substrate selectivity is produced. This may not be a requirement of a biotransformation catalyst, where the enzyme is presented with only one substrate species for reduction, but, when a mixture of potential substrates is present, such as may occur in a biological sample, this may be essential for achievement of selective conversion or the determination of one particular chemical species. This alteration in substrate specificity could also be advantageous in a biotransformation using whole cells where the intended substrate is necessarily contaminated with other entities which could also be transformed.

In the work of Wilks et al (Biochemistry, 1990, 27, 8587) a mutation strategy is described for the production of NAD-dependent dehydrogenases which have altered substrate specificity. The disclosed enzymes catalyse the reduction of homologues of pyruvic acid corresponding to the general formula: C_nH_{2n+1} CO COOH, which may include straight- and branched-chain alkyl residues. The initial intention of the present work was to continue the design method for substrates with an aromatic function, in addition to extended alkyl residues and hydroxyl and keto substitution associated with the same base structure of α-oxoacids.

Enzymes capable of reducing such substrates would be of particular value in the field of synthetic chemistry where an

α-keto compound could be converted stereospecifically to the corresponding secondary alcohol. The production of individual optical isomers of secondary alcohols is especially valuable in the manufacture of optical isomers of pharmaceuticals and drug intermediaries. The feature of thermophilicity which may be obtained with some α-hydroxy acid dehydrogenases is valuable as it enables the enzymic reactions to be carried out at relatively high temperature where a rate acceleration may exist and the enzymes are inherently stable. These enzymes may also be suitable for incorporation into determinations of the levels of particular substrates obtained in biological samples under certain disease states.

A numbering convention has evolved in the field of NAD-dependent dehydrogenases, which was originally based on an X-ray structure of dogfish muscle lactate dehydrogenase. This system numbers amino acids in ascending order extending from the N terminus. This system identifies conserved residues, such as glycine at positions 30 and 33, tyrosine at position 85, arginine at position 109, serine at position 163 and aspartic acid at position 168.

Thus, in any given NAD dependant dehydrogenase, natural or subject to mutation, there are regions of sequence which are homologous with the amino acid sequence of the numbering convention. An important aspect of this convention is that any amino acid change in an NAD dependent dehydrogenase may be accurately described.

In Table 1 below, an alignment of amino acid sequences is shown for three NAD dependent lactate dehydrogenases: the M4 isoenzyme of pig, the testis isoenzyme of man and the Bacillus stearothermophilus enzyme. (The symbols " - " do not signify breaks in the continuous polypeptide chains, instead they are conventional representation of discontinuities of numbering which allow alignment with sequences of other enzymes to give maximum homology.)

Table 1

	1				5				10				15				20	23	24	25			30					
	A	T	L	K	E	K	L	I	A	P	V	A	Q	Q	E	T	T	I	P	N	N	K	I	T	V	V	G	V	G	-	Q	V	G	M
	S	T	V	K	E	Q	L	I	E	K	L	I	E	D	D	E	-	-	S	Q	C	K	I	T	I	V	G	T	G	-	A	V	G	M
															M	K	N	N	G	G	A	R	V	V	V	I	G	A	G	-	F	V	G	A

	35				40				45				50				55				60				65				70				75				80				85				90
	A	C	A	I	S	I	L	G	K	-	-	S	-	L	-	-	-	-	T	D	E	L	A	L	V	D	V	L	-	-	E	D	K										
	A	C	A	I	S	I	L	L	K	-	-	D	-	L	-	-	-	-	A	D	E	L	A	L	V	D	V	A	-	-	L	D	K										
	S	Y	V	F	A	L	M	N	Q	-	-	G	-	I	-	-	-	-	A	D	E	I	V	L	I	D	A	N	-	-	E	S	K										
	L	K	G	E	M	M	D	L	Q	H	G	S	L	F	L	Q	T	P	K	I	V	A	N	K	D	Y	-	S	V	T	A	-	N										
	L	K	G	E	M	M	D	L	Q	H	G	S	L	F	F	S	T	S	K	V	T	S	G	K	D	Y	-	S	V	S	A	-	N										
	A	I	G	D	A	M	D	F	N	H	G	K	V	F	A	P	K	P	V	D	I	W	H	G	D	Y	-	D	D	C	R	-	D										
	S	K	I	V	V	V	T	A	G	-	V	R	Q	Q	E	G	E	S	R	L	N	L	V	Q	R	N	V	N	V	F	K	F	I										
	S	R	I	V	I	V	T	A	G	-	A	R	Q	Q	E	G	E	T	R	L	A	L	V	Q	R	N	V	A	I	M	K	I	I										
	A	D	L	V	V	I	C	A	G	-	A	N	Q	K	P	G	E	T	R	L	D	L	V	D	K	N	I	A	I	F	R	S	I										
	I	P	Q	I	V	K	Y	S	P	-	N	C	I	I	I	V	V	S	N	P	V	D	I	L	T	Y	V	T	W	K	L	S	G										
	I	P	A	I	V	H	Y	S	P	-	D	C	K	I	L	V	V	S	N	P	V	D	I	L	T	Y	I	V	W	K	I	S	G										
	V	E	S	V	M	A	S	G	F	-	Q	G	L	F	L	V	A	T	N	P	V	D	I	L	T	Y	A	T	W	K	F	S	G										
	-	-	-	-	L	P	K	H	R	V	I	G	S	G	C	N	L	D	S	A	R	F	R	Y	L	M	A	E	K	L	G	V	H										
	-	-	-	-	L	P	V	T	R	V	I	G	S	G	C	N	L	D	S	A	R	F	R	Y	L	I	G	E	K	L	G	V	H										
	-	-	-	-	L	P	H	E	R	V	I	G	S	G	T	I	L	D	T	A	R	F	R	F	L	L	G	E	Y	F	S	V	A										
	P	S	S	C	H	G	W	I	L	-	G	E	H	G	D	-	S	S	V	A	V	W	S	G	V	N	V	A	G	V	S	-	L										
	P	T	S	C	H	G	W	I	I	-	G	E	H	G	D	-	S	S	V	P	L	W	S	G	V	N	V	A	G	V	A	-	L										
	P	Q	N	V	H	A	Y	I	I	-	G	E	H	G	D	-	T	E	L	P	V	W	S	Q	A	Y	I	G	V	M	P	-	I										

Table 1 (cont)

```
    1          2          2          2          2          2          2
5   0          1          2          3          3          4
    B          5          0          0          5          0
    Q Q L N P E M G T D N D S E N W K E V H K M V V E S A Y E V I K L
    K T L D P K L G T D S D K E H W K N I H K Q V I Q S A Y E I I K L
    R K L V E S K G E E A Q K D - L E R I F V N V R D A A Y Q I I E K
10

          2          2          2          2          2          2
          4          5          5          6          6          7
          5          0          5          0          5          0
15  K - G Y T N - W A I G L S V A D L I E S M L K N - - L S R I H P V
    K - G Y T S - W A I G L S V M D L V P - - L K N - - L R R V H P V
    K - G A T Y - Y G I A M G L A R V T R A I L H N - - E N A I L T V

20        2          2          2          2          2          2  3
          7          8          8          9          9          9  0
          5          0          5          0          5          9  1
    S T M V Q G - - - - M Y G I E N E V F L S L P C V L N A R G L T S
    S T M V K G - - - - L Y G I K E E L F L S I P C V L G R N G V S D
25  S A Y L D G - - - - L Y G E R D - V Y I G V P A V I N R N G I R E

          3          3          3          3          3          3  3  3
          0          1          1          2          2          0  0  3
          5          0          5          0          5          A  B  1
    - V I N Q K L K D D E V A Q L K N S A D T L W G I Q K D L K D L
    - V V K I D L S E E E - A L L K K S A E T L W N I Q K N L I - F
    - V I E I E L N D D E K N R F H H S A A T L K S V L A R A F T R
35
```

Expression cloning of human testis-specific lactate dehydrogenase cDNA.
Millan, J.L., Driscoll, C.E. and Goldberg E.
Sequence from cDNA - Genbank accession number Jϕ2938 (1986).

40 The DNA sequence of the thermophilic lactate dehydrogenase from Bacillus stearothermophilus.
Barstow, D., Clarke, A.R., Wigley, D., Holbrook, J.J. and Atkinson, T.
Gene, 46, (1986), 47-55

Within the conventional numbering system are short sequences which may be correlated with specific structural elements in the folded polypeptide and which may have specific functional properties such as the substrate recognition site or the activation site.

The substrate recognition site is carried in part by a mobile loop of polypeptide chain, conventionally numbered 98 to 110. This sequence is contiguous but traditionally omits a residue 103.

It is known for α-hydroxy acid dehydrogenases of the L type which generate S stereochemistry on reduction to the hydroxy function that a mobile surface loop exists which changes conformation after substrate binding. This loop consists of the amino acid residues 98-110 and contains an arginine at position 109 which is important for catalysis as the positive charge from the amidine group stabilises the stretched substrate carbonyl and thus decreases the energy required to obtain the transition state necessary for hydride transfer.

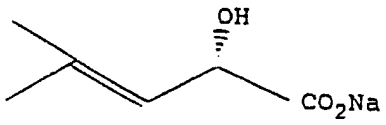
The loop region is also involved in substrate selection and for that reason was the particular object for the present enzyme engineering study.

The mechanism by which lactate dehydrogenase distinguishes different substrates is the ability of the substrate to fit into a proton-impermeable, fixed-sized internal vacuole which is formed when the mobile surface polypeptide loop closes down onto the protein surface. Not only is loop closure only possible over suitably small and singly negatively charged substrates, but also the loop closure triggers catalysis through the arginine 109 residue. The variation in composition and length of this mobile loop region is the immediate object. For the convenience of these experiments, a particular gene for wild-type Bacillus stearothermophilus lactate dehydrogenase was chosen where the amino acids alanine at positions 235 and 236 had been

changed for glycines. The effects of this particular amino acid substitution have been presented by Wilks et al. for a limited range of substrates (Biochemistry, 28, 8587) and generally increased the activity towards substrates with larger alkyl groups. Although used to demonstrate the principle of loop exchange, the technique would not be constrained to this particular enzyme, rather it is applicable not only to this mutant enzyme, but also to all other structurally-related α-hydroxy acid dehydrogenases, for example.

The mutation where alanines at 235,236 are replaced by glycines has been combined with three mutations in the mobile polypeptide loop (residues 98-112), namely glycine 102 by methionine, lysine 103 by valine and proline 105 by serine (MVS/GG).

This new enzyme construction was evaluated for activity towards longer substrates, in particular an unsaturated branched substrate 4-methyl-2-oxo-3-pentenoic acid, which is reduced to the following alcohol:



Steady state kinetic measurements indicated that reduction of this compound by the wild-type enzyme proceeded slowly, obtaining an estimate for turnover of $0.03S^{-1}$ in contrast to that obtained with the mutant enzyme of $1.2S^{-1}$. The Km determined under similar conditions of substrate concentration (1-20mM) in the presence of 5mM fructose 1,6-bisphosphate was 22mM. This observation regarding the specificity alteration towards a less flexible substrate indicates that the loop region has importance in substrate reduction.

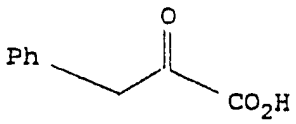
The method used to make new loop constructions was to insert restriction enzyme sites at either end of the DNA coding

for the loop region. These new restriction sites which are unique within the DNA coding for the enzyme, are cleaved and then religated with synthetic DNA designed to code for the required new loop region. One of the restriction sites introduced was for SacII near amino acid 97. The construction of the Sac II restriction site required that the wild type coding sequence for cysteine 97 was changed to threonine. The XbaI site retained the wild-type amino acid sequence with arginine at 109, but did result in the creation of an MluI site close to threonine 108. The new MluI site was used to advantage as it was destroyed in transformants and thus enabled easy distinction thereof from the wild-type gene.

To illustrate the utility of the loop design approach to enzyme engineering, novel loops were introduced, two shorter by 3 amino acids and one longer by 4 amino acids. The new enzymes generated in this manner were evaluated against a range of experimental substrates to determine the effect of the loop exchanges.

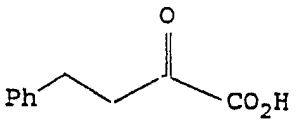
It was clearly demonstrated that the new loops altered the properties of the enzyme from that of the framework used in the construction thereof. The results also illustrate the difference obtained with the alanine → glycine alteration at amino acids 235 and 236 and the introduction of the threonine in place of cysteine at amino acid 97.

The increase in turnover of α-ketocaproate and α-ketoisocaproate with the alanine → glycine double mutation was consistent with the results of Wilks et al. (Biochemistry, 29, 1990, 8587). The increase in turnover for the aromatic substrate 2-oxo-4-phenyl propanoic acid:



along with increases in the Km for both was not obvious and indicates useful improvement with respect to the use of the mutant enzyme in the synthesis of the chiral α-hydroxy group of this aromatic substrate.

The exchange of threonine for cysteine at amino acid 97 maintained the beneficial Km effect for 2-oxo-4-phenyl butanoic acid:



over the wild-type enzyme.

The effect of these individual mutations on the reduction of the aromatic substrates is of clear interest as the homochiral hydroxyacids produced form useful chiral building blocks for the synthesis of bioactive compounds.

The introduction of the new loop sequences further alters the substrate specificity of the enzyme reducing the turnover of the natural substrate from that of the wild type enzyme. The three new loop enzymes retained most of the wild type catalytic potential towards the 2-oxo-4-phenyl propanoic acid as shown by turnover and Km and, in the example of the longer loop and second shorter loop version, resulted in an increase in turnover.

These examples serve to illustrate that the activity of the enzyme may be dramatically altered by changes in the loop sequences, both towards the natural substrate and larger unnatural substrates.

In the large loop, it is observed that the Kcat/Km for

2-oxo-4-phenyl propanoic acid was 1700 times better than for pyruvate compared to the wild type enzyme which is conversely 230 times better for pyruvate, representing a switch in specificity of 391,000 fold.

The alteration in specificity of the enzyme from pyruvate to 2-oxo-4-phenyl propanoic acid renders the new enzyme suitable for the determination of the concentration of 2-oxo-4-phenyl propanoic acid, often termed phenyl pyruvate in clinical chemistry nomenclature, especially from body fluids, such as blood and urine.

Phenyl pyruvate levels are normally low, but rise to significant levels with the increase in phenylalanine concentration, which is associated with the genetic disease phenylketonuria (Langenbeck et al., J. Inher. Metab. Dis., 4, 1981, 69). It is also possible that the phenyl pyruvate reductase or phenyl lactate dehydrogenase enzyme could be used in conjugation with phenylalanine dehydrogenase, a current method of determining the phenylketonuria level such that interference from phenyl pyruvate could be negated, thereby enhancing the sensitivity of the phenylalanine-based method.

The construct having the restriction sites at either end of the loop region may be used to produce a series of dehydrogenases having loops of variable length and variable sequence. Thus, by restricting random mutagenesis to the region of lactate dehydrogenase which has been identified as being important for substrate recognition, it is possible to isolate enzymes which may carry out a desired chiral reduction. The random mutagenesis may be generated by use of spiked oligonucleotides at specific positions and on different length loops or, alternatively, by the incorporation of inosine triphosphate in a polymerase chain reaction (PCR) that randomises either the entire loop region or specific residues. Both of these techniques have been employed to prepare mutant libraries using the restriction sites engineered into the DNA coding for

the loop region of LDH. A further PCR method was used to generate a random combinational DNA library of specific positions of the loop region. This technique was specifically targeted to positions 101 and 102 as these are involved in defining enzyme substrate specificity.

The PCR was initially used to generate 300 & 800 base pair fragments that had complementary overlapping ends. These primary products which had random sequences incorporated in the overlap, were then primed on each other and extended to yield an LDH hybrid gene. A second PCR with two outer primers annealing at non-overlapping ends was finally used to amplify the LDH product.

Previous manipulation of the Bacillus stearothermophilus LDH gene involved cloning an EcoRI/PstI digested gene in to PKK 233-2, or M13 plasmid vectors. Where, as now, it is possible to clone the PCR product into any one of a number of vectors, because one of the outer primers (2), which anneals past the coding region, was designed with an additional EcoRI site incorporated. For example, in order to verify that there is a representative library with random sequences in the desired positions, it is possible to clone the gene with unique EcoRI sites into PUC18, which produces a high yield of DNA from mini-preps, and subsequently the PCR product may be cloned into plasmid or phage expression vectors, such as PKK 233-2. (See accompanying illustrative Fig. 2.)

The following advantages are obtained with the PCR method:

1. High yield of PCR product obtained.
2. The ability to identify product as mutant DNA and select against wild-type sequences via MluI digestion.
3. Ease of handling and monitoring a 1kb product compared to previous attempts which involved designing restriction sites either side of the loop region, such that a 40 base pair wild-type sequence may be replaced with a mutant sequence.

4. Speed of method.

5. The design of primer 2 with an EcoRI site enables the cloning of gene product into a number of vectors.

6. Use of double-stranded template for mutagenesis.

7. Application of method to manipulate other areas of the LDH gene and the ease by which interesting mutations in different regions may be brought together in one molecule using this splice overlap extension method.

8. Having mutant oligos with a high region of complementarity to the template at the 3'-end ensures that annealing of oligos to the vector is highly efficient.

In order successfully to utilise a directed random mutagenesis method that generates a library of mutants covering the loop region of the enzyme, or indeed any specific region of any target enzyme, requires a suitable screen for clones which express mutant enzymes of the desired specificity. For the dehydrogenases, this is simply provided by coupling NADH production with phenazine metasulphate to formation of insoluble blue formazan dye.

The screen is based on the work of Katzen and Schimkel (PNAS, 54, 1218) and relies on the ability of a colony expressing an enzyme with specificity to oxidise the required substrate and to reduce NAD^+ to NADH. The reduced coenzyme then reduces phenazine metasulphate which in turn reduces nitroblue tetrazolium to form an insoluble blue dye.

The mutant DNA is transformed into competent E. coli cells and is stored on agar plates containing 15% glycerol and ampicillin at -80°C. Obtaining electro-competent cells with high transformation rates has produced rates of 10^6 per μg of DNA, a rate which produces a sufficiently representative population of mutant colonies for screening. Copies of this plate are made using a velvet replicator and the copies grown up overnight. (The E. coli LDH activity is removed by incubation of the filter paper at 67°C for 30 minutes, the activity of the wild-type

enzyme is not lost until 45 minutes at this temperature.) The copies are then screened against a range of substrates and individual colonies may be compared. Each master plate is screened at least three times to ensure conditions are ideal in each case.

Using this technique demonstrates differential rates of staining have been shown between filter copies of wild-type colonies and those containing the malate dehydrogenase activity mutant enzyme (Q102R) with lactate and malate as substrates, respectively, confirming the validity of the screen to identify individual colonies.

The following illustrates the present invention:

Mutagenesis of lactate dehydrogenase

Mutants of lactate dehydrogenase from Bacillus stearothermophilus were generated by the oligonucleotide mismatch procedure of Winter et al. (Nature, 1982, 299, 756) in M13 with the mutagenic oligonucleotide as the primer for in vitro chain extensions. The double alanine replacement at 235 and 236 by glycine was obtained using the oligonucleotide sequence 3'CGCGCTACCGCCGATGTTTA5'. The wild type and mutant enzymes were expressed in the PKK223-3 plasmid in E. coli (Barstow et al., Gene, 1986, 46, 47).

Mutagenesis to construct Sac II and XbaI sites at either end of the gene coding for wild type active site loop.

A 54-mer oligonucleotide was used to direct mutagenesis to introduce unique restriction sites (SacII and XbaI) at either end of the active site loop (amino acids 98-110) using the wild-type template (Barstow loc. cit). The mutagenic oligonucleotide was:

5'GTCCACAAGGTCTAGACGCGTCTCGCCCGGTTTTTGGTTGGCGCCCGCGGTAATGACAAC3',

the annealing, chain extension and cloning were as described by Clarke et al. (Nature, 1986, 329, 699).

Mutants were identified by making mini-preps and restricting with SacII and XbaI. Mutant mini-preps were restricted with EcoRI and XhoI and the small fragment was sub-cloned into PKK223-3 containing Ala235Gly, Ala236Gly mutant LDH from which the small EcoRI/XhoI fragment had been removed (Wilks et al. Biochemistry, 1990, 29, 8587). The resulting plasmid (pLDHrs) was transformed into competent E. coli TG2 cells. The whole sequence was redetermined using a "Dupont Genesis 2000" automatic sequencer and showed the correct loop sequence had been inserted. The partial DNA sequences of the wild type gene and the mutant with inserted restriction sites are shown in Table 2 below.

Table 2 *Comparison of the protein and DNA sequences of B. stearothermophilus lactate dehydrogenase in the loop (93-111) region of wild-type and the mutant with SacII and XbaI restriction sites at either end of the loop, and the variable loop sequences derived from them.*

Wild-type DNA sequence in loop region (Cys changed to Thr is shown bold):

```
  LeuValValIleCysAlaGlyAlaAsnGlnLysProGlyGluThrArgLeuAsp
5'TTGGTTGCTATTTGCGCCGGCGCCAACCAAAAACCGGGCGAGACGCGGCTTGAT3'
3'AACCAACGATAAACGCGGCCGCGGTTGGTTTTTGGCCCGCTCTGCGCCGAACTA5'
```

Mutant DNA (pLDHrs) sequence in loop region:

```
LeuValValIleThrAlaGlyAlaAsnGlnLysProGlyGluThrArgLeuAsp
5'TTGGTTGCTATTACCGCGGGCGCCAACCAAAAACCGGGCGAGACGCGTCTAGAC3'
3'AACCAACGATAATGGCGCCCGCGGTTGGTTTTTGGCCCGCTCTGCGCAGATCTG5'
              SacII                                XbaI
                                                   MluI
```

Two oligonucleotides (LLA and LLB) used to synthesise the big loop by PCR:

```
5'TACCGCGGGCAACATTAAATTGCAACAAGATAA3'   (LLA)
    SacII
5'GGTCTAGACGATCGCCCGTCGGGTTATCTTGTT3'   (LLB)
    XbaI
```

Big loop sequence in the 97-110 region (note the MluI site is destroyed):

```
 CysAlaGlyAlaAsnGlnLys----------ProGlyGluThrArgLeuAsp (wild-type)
 ThrAlaGlyAsnIleLysLeuGlnGlnAspAsnProThrGlyAspArgLeuAsp (big loop)
5'TACCGCGGGCAACATTAAATTGCAACAAGATAACCCGACGGGCGATCGTCTAGACC3'
3'ATGGCGCCCGTTGTAATTTAACGTTGTTCTATTGGGCTGCCCGCTAGCAGATCTGG5'
    SacII                                              XbaI
```

Oligonucleotides for PCR synthesis of LeuLysGly and SerLysGly short loops:

```
SLA  5'TACCGCGGGCGCCAACT3'
SLB  5'GGTCTAGACGGCCTTTCAAGTTGGCGCC3'
SLC  5'GGTCTAGACGGCCTTTGGAGTTGGCGCC3'
```

Short loop sequence in the original 97-111 region (MluI site is again destroyed):

```
                      GlyGluThr
  CysAlaGlyAlaAsnGlnLysProArgLeuAsp       (wild-type)
  ThrAlaGlyAlaAsnLeuLysGlyArgLeuAsp       (SL1)
5'TACCGCGGGCGCCAACTTGAAAGGCCGTCTAGACC3'
3'ATGGCGCCCGCGGTTGAACTTTCCGGCAGATCTGG5'

  ThrAlaGlyAlaAsnSerLysGlyArgLeuAsp       (SL2)
5'TACCGCGGGCGCCAACTCCAAAGGCCGTCTAGACC3'
3'ATGGCGCCCGCGGTTGAGGTTTCCGGCAGATCTGG5'
    SacII                    XbaI
```

PCR assembly method for generation of random combinational library of the loop region of the B. stearothermophilus LDH gene:

1. Single-stranded oligos were made such that the oligos were only different to the wild-type sequence at positions encoding amino acids 101 and 102 where each one of the bases A, T, C, G has an equal chance of being inserted. (Oligo mix 101,102 forward.)

2. An MluI restriction site which is present in the wild-type template is destroyed by change of the third codon position of amino acid 108 from an ACG to an ACT without altering threonine as the amino acid being coded. The absence of the MluI site enables verification that the mutants have been generated and to select against wild-type sequences.

3. A DNA primer which has 14 base homology to olio mix 101,102 forward was used to make the complementary strand (oligo mix 101,102 reverse) using a Klenow reaction.

4. Single-stranded library oligos were used with primer 1 and 5ng of wild-type template in order to generate a 300 base pair product with 25 cycles of PCR (94°C, for 1 minute, 55°C for 1 minute, 72°C for 2 minutes).

5. Double-stranded Klenow oligos were used with primer 2 and 5ng of wild-type template to generate an 800 base pair product which overlaps the 300 base pair product. (PCR conditions as in 4.)

The use of double-stranded oligo as primer in 5 is very important in ensuring that both the 300 and 800 base pair products are made and primed using mutant oligos and that the wild-type sequence at position 101 and 102 is not copied.

6. After gel purification, 20ng of the 300 base pair product and 60 ng of the 800 base pair product were mixed without primers and thermocycled seven times in order to join the fragments (94°C for 2 minutes, 55°C for 1 minute, 72°C for 4 minutes).

7. After seven cycles, primers 1 and 2 were added, and the product amplified for twenty cycles (94°C for 1.5 minutes, 55°C for 1 minute, 72°C for 2.5 minutes).

8. The 1 kb PCR product was then gel purified, digested with EcoRI, and gel purified again before ligation into EcoRI-cut PUC18 plasmid vector and transformation into E. coli.

9. Recombinant colonies were selected for by IPTG and X-Gal insertional inactivation.

10. Of the nine white colonies picked, seven were verified for the presence of the LDH gene and to resistance to MluI digestion via gel and restriction analysis. The other two did not have inserts.

11. Six of the mutants were sequenced using a Dupont 2000 sequencer and confirm that the random mutagenesis approach had been achieved.

See Table 3 below:

Table 3

1	TTG	GTT	GCT	ATT	ACC	GCG	GGC	GCC	TCC	CCC	AAA	CCG	GGC	GAG	ACT	CGT	CTA	GAC
2	TTG	GTT	GCT	ATT	ACC	GCG	GGC	GCC	TCT	CTA	AAA	CCG	GGC	GAG	ACT	CGT	CTA	GAC
3	TTG	GTT	GCT	ATT	ACC	GCG	GGC	GCC	AAA	TTT	AAA	CCG	GGC	GAG	ACT	CGT	CTA	GAC
5	TTG	GTT	GCT	ATT	ACC	GCG	GGC	GCC	TCC	AAC	AAA	CCG	GGC	GAG	ACT	CGT	CTA	GAC
7	TTG	GTT	GCT	ATT	ACC	GCG	GGC	GCC	CCG	GCC	AAA	CCG	GGC	GAG	ACT	CGT	CTA	GAC
8	TTG	GTT	GCT	ATT	ACC	GCG	GGC	GCC	TTA	CGT	AAA	CCG	GGC	GAG	ACT	CGT	CTA	GAC
121	TTG	GTT	GCT	ATT	ACC	GCG	GGC	GCC	ACC	TTC	AAA	CCG	GGC	GAG	ACT	CGT	CTA	GAC
wtrs	TTG	GTT	GCT	ATT	ACC	GCG	GGC	GCC	AAC	CAA	AAA	CCG	GGC	GAG	ACG	CGT	CTA	GAC
wt	TTG	GTT	GCT	ATT	TGC	GCC	GGC	GCC	AAC	CAA	AAA	CCG	GGC	GAG	ACG	CGG	CTT	GAT

	93	94	95	96	97	98	99	100	101	102	103	104	106	107	108	109	110	111
1	Leu	Val	Ala	Ile	Thr	Ala	Gly	Ala	Ser	Pro	Lys	Pro	Gly	Glu	Thr	Arg	Leu	Asp
2	Leu	Val	Ala	Ile	Thr	Ala	Gly	Ala	Ser	Leu	Lys	Pro	Gly	Glu	Thr	Arg	Leu	Asp
3	Leu	Val	Ala	Ile	Thr	Ala	Gly	Ala	Lys	Phe	Lys	Pro	Gly	Glu	Thr	Arg	Leu	Asp
5	Leu	Val	Ala	Ile	Thr	Ala	Gly	Ala	Ser	Asn	Lys	Pro	Gly	Glu	Thr	Arg	Leu	Asp
7	Leu	Val	Ala	Ile	Thr	Ala	Gly	Ala	Pro	Ala	Lys	Pro	Gly	Glu	Thr	Arg	Leu	Asp
8	Leu	Val	Ala	Ile	Thr	Ala	Gly	Ala	Leu	Pro	Lys	Pro	Gly	Glu	Thr	Arg	Leu	Asp
121	Leu	Val	Ala	Ile	Thr	Ala	Gly	Ala	Thr	Phe	Lys	Pro	Gly	Glu	Thr	Arg	Leu	Asp
wtrs	Leu	Val	Ala	Ile	Thr	Ala	Gly	Ala	Asn	Gln	Lys	Pro	Gly	Glu	Thr	Arg	Leu	Asp
wt	Leu	Val	Ala	Ile	Cys	Ala	Gly	Ala	Asn	Gln	Lys	Pro	Gly	Glu	Thr	Arg	Leu	Asp

Generation of double-stranded DNA loop fragments by oligonucleotide-overlap

Each pair of overlapping oligonucleotides (20μM of each) were subjected to 30 cycles of annealing and extension (94°C for 1 minute, cool to 45°C for 2 minutes, 45°C for 1 minute, heat to 72°C in 1 minute, 72°C for 1 minute in 50μl containing 0.05 M KCl, 10mM Tris pH 8.3, 1.5 mM $MgCl_2$, 0.01% gelatin), 200μM of each dNTP and 2.5 units TAQ DNA polymerase). The double-stranded DNA product was purified and then cut with SacII and XbaI before ligating it into the plasmid pLDHrs cut with the same enzymes. The ligated products were restricted with MluI to cleave wild-type plasmid pLDHrs.

The DNA was purified, microdialysed and used to transform E. coli TG2 cells by electroporation. Transformed cells were selected for ampicillin resistance. Ten such colonies were picked and plasmid DNA purified from overnight cultures. The presence of mutant loops was confirmed by resistance to MluI digestion.

The expression of the enzymes was obtained as described above.

Purification of lactate dehydrogenase and mutants

Overnight cultures (1l) were centrifuged and the packed cells were resuspended in 50 mM triethanolamine, pH 6.0. The cells were sonicated and the debris was removed by centrifugation. The protein in the supernatant was precipitated by the addition of 65% ammonium sulphate. The precipitate was spun down and resuspended in 50 mM triethanolamine, pH 6.0 and dialysed against the same buffer. After dialysis, NADH and FBP were added to the protein to final concentrations of 5mM and 10mM before loading onto an oxamate Sepharose column which had been

pre-equilibrated with 50 mM triethanoloamine, pH 6.0, 3 mM NADH and 5 mM FBP. After washing off unbound protein with column buffer mutant LDH was eluted with 50 mM triethanolamine, pH 9.0, 0.3 M NaCl. The elutant was precipitated with 65% ammonium sulphate and then resuspended in and dialysed against 50 mM triethanolamine, pH 7.5. The protein was then loaded onto a Q-Sepharose Fast Flow column and eluted with a salt gradient. LDH eluted at a concentration of 0.25 M NaCl. For the double glycine mutant enzyme, the first chromatography procedure with oxamate Sepharose was replaced by chromatography on Blue Sepharose -F3GA, otherwise the procedure was essentially the same. All proteins were judged to be greater than 98% pure from the intensity of Coomassie blue staining on an SDS Phast gel (Pharmacia). The yield of protein was usually 0.2g/l of original broth.

Steady-State Kinetics

Steady-state measurements were made by following the reduction in absorbance at 340nm in the NADH/NAD^+ conversion. All assays were at 25°C in the buffer Bis-Tris, pH 6, (20mM), containing KCl (50mM) and when used fructose-1,6- bisphosphate at 5mM. Protein concentration was determined from the absorbance at 280nm using the value of 0.91 for 1mg/ml protein in 1cm path and an Mr of 33,000.

The results from these determinations are shown in Table 4 below.

TABLE 4 Steady State Kinetic Parameters of some Loop Exchange Mutants

	ENZYME	WILD-TYPE +FBP	WILD-TYPE -FBP	$^{235/6}$GG +FBP	$^{235/6}$GG -FBP	WTrs +FBP	WTrs -FBP	BIG LOOP +FBP	BIG LOOP -FBP	SL1 +FBP	SL1 -FBP	SL2 +FBP	SL2 -FBP
PYRUVATE	$k_{cat}s^{-1}$	250	250	167	-	60	19	0.2	0.05	0.2	0.1	0.07	0.04
	K_mmM	0.06	2	4	-	3.5	100	42	50	0.08	2.0	0.065	0.5
	k_{cat}/K_M $M^{-1}.s^{-1}$	4.2E6	5E4	4.2E4	-	1.7E4	-	4.7	-	2.5E3	-	1076	80
KETO	$k_{cat}s^{-1}$	29	-	240	-	88	12	6	0.3	0.8	0.07	0.1	0.1
CAPROATE	K_mmM	3.4	-	5.6	-	5.8	30	20	27	16	20	25	60
	k_{cat}/K_M $M^{-1}.s^{-1}$	8.5E3	-	4.2E4	-	1.5E4	-	300	-	50	-	4	1.6
KETOISO	$k_{cat}s^{-1}$	0.33	-	1.74	-	1.8	0.2	0.3	0.06	0.07	0.01	0.9	0.08
CAPROATE	K_mmM	6.7	-	15.4	-	4	28	18	32	20	40	25	30
	k_{cat}/K_M $M^{-1}.s^{-1}$	50	-	112	-	450	-	17	-	3.5	-	36	2.6
2-OXO-4-PHENYL	$k_{cat}S^{-1}$	6	-	7	-	6	0.8	1.0	0.1	0.03	0.002	0.2	0.01
BUTANOATE	K_mmM	0.6	-	13	-	4	12	7	12	4	4	4	4
	k_{cat}/K_M $M^{-1}.s^{-1}$	1E4	-	538	-	1.5E3	-	143	-	7.5	-	50	2.5
2-OXO-4-PHENYL	$k_{cat}s^{-1}$	32.7	-	53.4	-	58	4	40	6	20	10	100	20
PROPANOATE	K_mmM	1.8	-	4.5	-	6	21	5	100	11	80	3	20
	k_{cat}/K_M $M^{-1}.s^{-1}$	1.8E4	-	1.2E4	-	9.6E3	-	8E3	-	1.8E3	-	3.3E4	1E3

K_m values above 50mM are less accurate due to the large substrate absorbance

Reduction of 4-methyl-2-oxo-3-pentenoic acid using MVS/GG:

MVS/GG (6 units (μ moles/minute/30°C)) and yeast formate dehydrogenase (5 units) were added to a solution of 4-methyl-2-oxo-3-pentenoic acid (1.0 mM) in deoxygenated Tris buffer (5mM:pH 6.0; 80ml) containing NADH (0.02 mM), sodium formate (3.1 mM), fructose-1,6-bisphosphate (0.4 mM) and dithiothreitol (0.08 mM). The solution was stirred at room temperature (~20°C) under nitrogen for 5 days with periodic addition of 0.2 mM HCl to maintain pH in the range of 6.0 - 6.2. Acidification to pH 2.0 and extractive work-up with ethyl acetate gave (S)-2-hydroxy-4-methyl-3-pentenoic acid in 91% isolated yield. Analysis of the (R)-MTPA derivative and comparison to a racemic standard gave a value of at least 99% for entantiomeric excess.

Claims:-

1. A method for modifying the specificity and/or efficiency of an enzyme, while retaining its catalytic activity, characterised in that it comprises: selecting an enzyme, the tertiary structure of which is substantially known or deduced; identifying at least one specificity and/or efficiency-related region; identifying or constructing unique restriction sites bounding the identified region in the DNA coding therefor; generating a DNA sequence which corresponds to at least a portion of the identified region, except that the nucleotides of at least one codon are randomized, or selecting as a substitute for at least a portion of the identified region an alternative such region, which may itself be similarly randomized; using the generated or substitute DNA sequence to replace the original such sequence; expressing the DNA including the generated or substitute DNA sequence; and selecting for a desired modification so that the DNA coding therefor may be isolated.

2. A method as claimed in claim 1 wherein the enzyme selected is a dehydrogenase.

3. A method as claimed in claim 2 wherein the dehydrogenase is an α-hydroxy acid dehydrogenase.

4. A method as claimed in any of claims 1 to 3 wherein a loop region of an enzyme is identified.

5. A method as claimed in any of claims 1 to 4 wherein the randomized DNA is generated by means of an inosine triphosphate PCR method or a spiked oligonucleotide method or a PCR assembly method.

6. A method as claimed in any of claims 1 to 5 wherein the selected substitute is based on a corresponding sequence from a similar enzyme.

7. A method as claimed in any of claims 1 to 6 wherein the generated or substitute DNA is cloned into a plasmid or phage vector and transformed into a bacteria or virus for expression.

8. A method as claimed in any of claims 1 to 7 wherein the enzyme is L-lactate dehydrogenase, positions 101 and 102 having been randomized.

9. A process for the production of a chiral product characterised in that it comprises the use of an enzyme which has been modified by a method as claimed in any of claims 1 to 8.

10. A process as claimed in claim 9 wherein a cofactor recycling system is provided.

11. A process for the reduction of 2-oxo-4-phenyl-propanoic acid characterised in that it comprises the use of L-lactate dehydrogenase, which has been modified in the loop region by a method as claimed in any of claims 1 to 8.

12. A process for the reduction of 4-methyl-2-oxo-3-pentenoic acid characterised in that it comprises the use of MVS/GG obtainable by a method as claimed in any of claims 1 to 8.

FIG. 1

ACTIVE SITE OF LACTATE DEHYDROGENASE

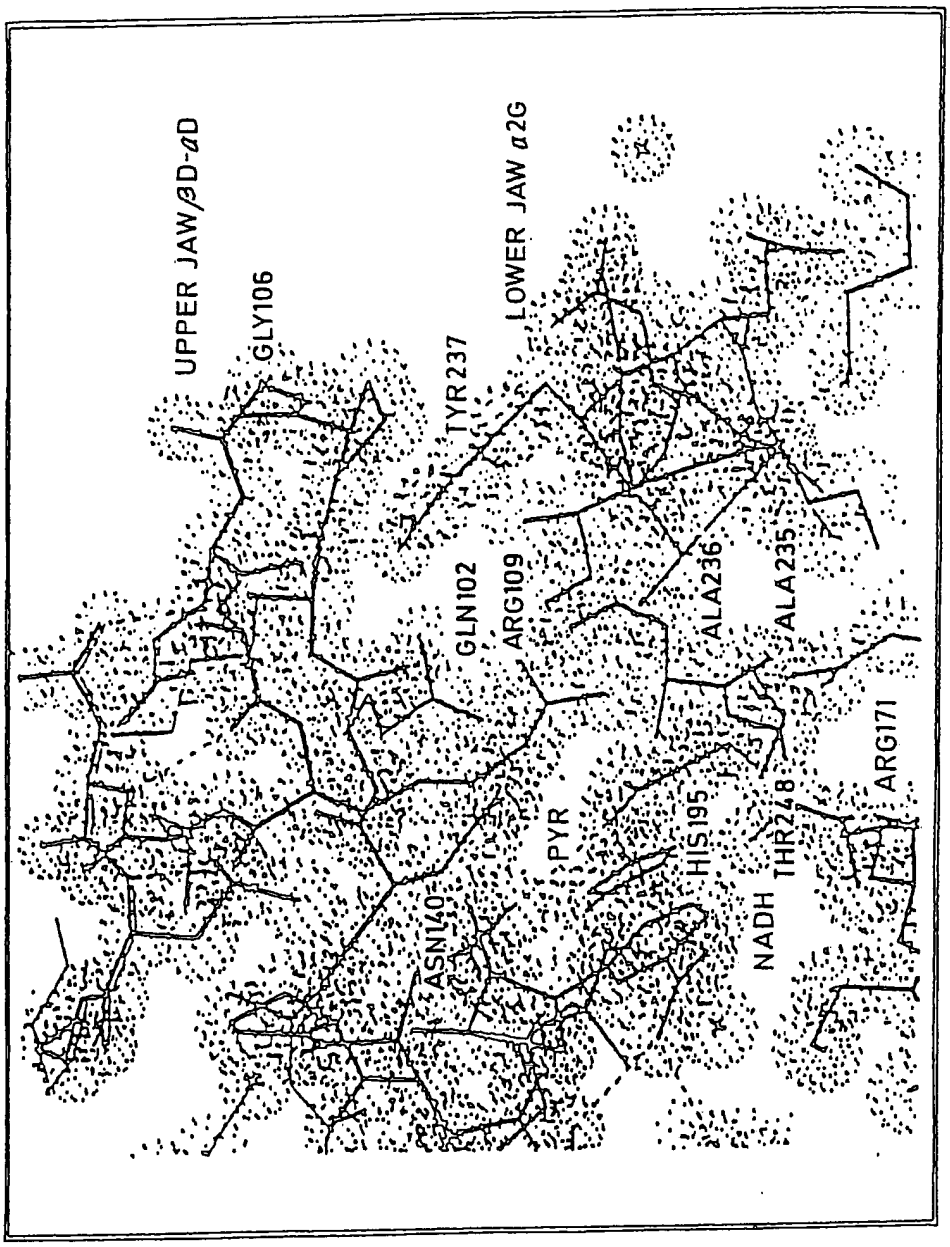
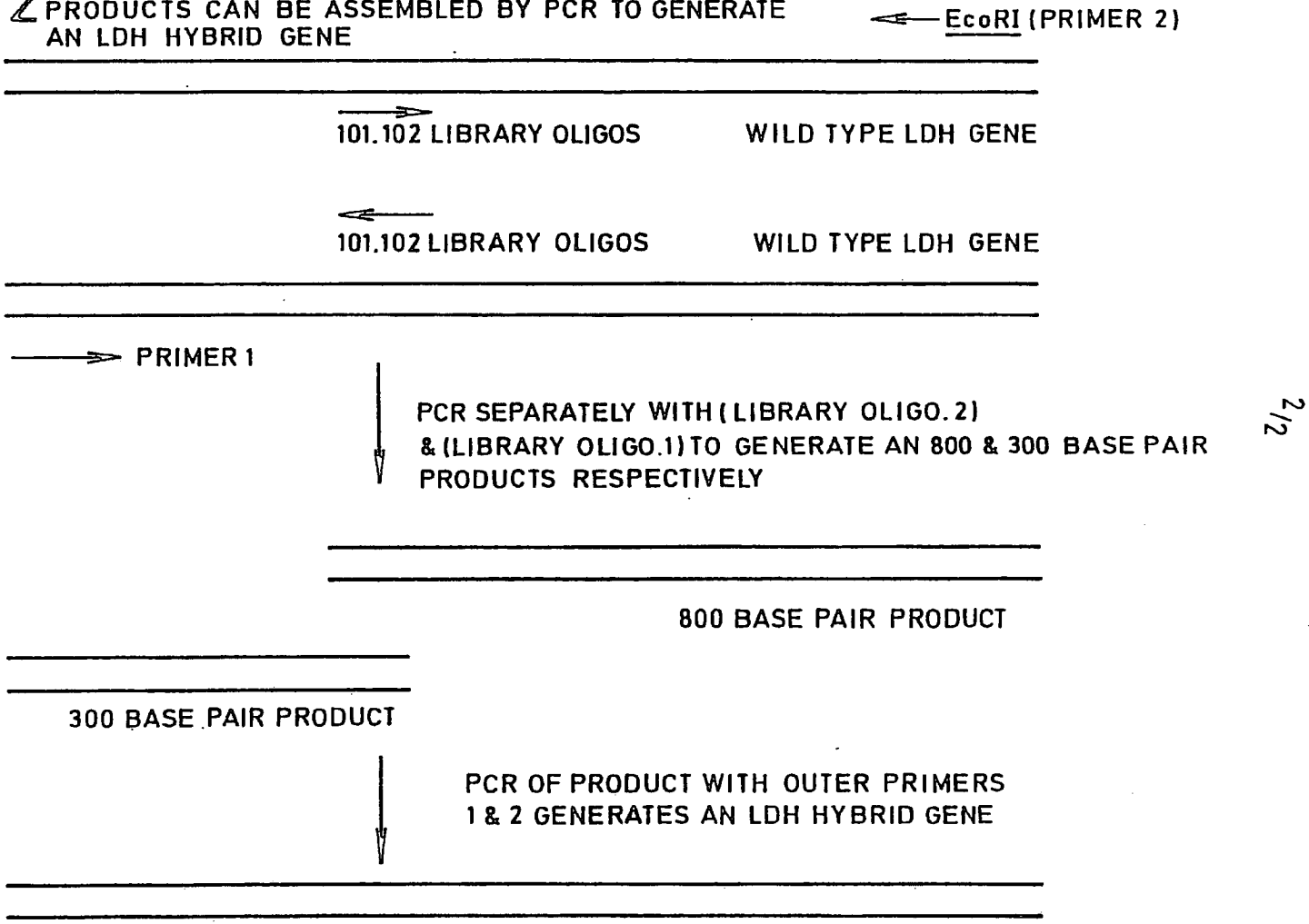


FIG. 2 DIAGRAM SHOWS HOW TWO OVER LAPPED PRIMARY PRODUCTS CAN BE ASSEMBLED BY PCR TO GENERATE AN LDH HYBRID GENE



INTERNATIONAL SEARCH REPORT

International Application No PCT/GB 93/00204

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all)[6]

According to International Patent Classification (IPC) or to both National Classification and IPC

Int.Cl. 5 C12N15/52; C12N15/53; C12N9/00; C12N9/04
C12P7/42

II. FIELDS SEARCHED

Minimum Documentation Searched[7]

Classification System	Classification Symbols
Int.Cl. 5	C12N ; C12P ; C12Q

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched[8]

III. DOCUMENTS CONSIDERED TO BE RELEVANT[9]

Category °	Citation of Document,[11] with indication, where appropriate, of the relevant passages[12]	Relevant to Claim No.[13]
X	BIOCHEMISTRY. vol. 29, 1990, EASTON, PA US pages 8587 - 8591 HELEN M. WILKS ET AL. 'Designs for a broad substrate specificity keto acid dehydrogenase.' cited in the application	1-4,7,8
Y	see the whole document see page 8591, column 1, line 38 - line 41	5,11
X	PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA. vol. 86, December 1989, WASHINGTON US pages 9094 - 9098 A.R.OLIPHANT ET AL. 'An efficient method for generating proteins with altered enzymatic properties:Application to beta-lactamase.' see the whole document	1,4,6,7

-/--

° Special categories of cited documents :[10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
07 MAY 1993	25.05.93
International Searching Authority	**Signature of Authorized Officer**
EUROPEAN PATENT OFFICE	DELANGHE L.L.M.

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

Category °	Citation of Document, with indication, where appropriate, of the relevant passages	Relevant to Claim No.
X	EP,A,0 251 446 (GENENTECH) 7 January 1988 see claims; example 14	1,4,6,7
X	EP,A,0 247 647 (GENENTECH) 2 December 1987 see claims see page 17, line 26 - line 29	1,4,6,7
Y	HENRY A. ERLICH 'PCR TECHNOLOGY:Principles and applications for DNA amplification.' 1989 , STOCKTON PRESS , NEW YORK see page 61 - page 70	5
Y	NUCLEIC ACIDS RESEARCH. vol. 16, no. 22, 1988, ARLINGTON, VIRGINIA US page 10932 KAREN KNOTH ET AL. 'Highly degenerate, ionosine-containing primers specifically amplify rare cDNA using the polymerase chain reaction.' see the whole document	5
Y	PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA. vol. 85, December 1988, WASHINGTON US pages 9436 - 9440 MICHAEL A. INNIS ET AL. 'DNA sequencing with Thermus aquaticus DNA polymerase and direct sequencing of polymerase chain reaction-amplified DNA.' see the whole document	5
Y	APPLIED BIOCHEMISTRY AND BIOTECHNOLOGY vol. 22, 1989, pages 169 - 179 ETHAN S. SIMON ET AL. 'D-Lactate Dehydrogenase:Substrate specificity and use as a catalyst in the synthesis of homochiral 2-Hydroxy acids.' see the whole document	11

ANNEX TO THE INTERNATIONAL SEARCH REPORT ON INTERNATIONAL PATENT APPLICATION NO.

GB 9300204
SA 69758

This annex lists the patent family members relating to the patent documents cited in the above-mentioned international search report. The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information. 07/05/93

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
EP-A-0251446	07-01-88	AU-B-	614929	19-09-91
		AU-A-	7228187	05-11-87
		JP-A-	1085075	30-03-89
		US-A-	5182204	26-01-93
		US-A-	5204015	20-04-93
EP-A-0247647	02-12-87	US-A-	4760025	26-07-88
		AU-B-	587960	07-09-89
		AU-A-	2956884	03-01-85
		AU-B-	631797	10-12-92
		AU-A-	3714989	23-11-89
		AU-A-	3720889	07-12-89
		EP-A-	0130756	09-01-85
		EP-A,B	0246678	25-11-87
		EP-A-	0357157	07-03-90
		JP-A-	60070075	20-04-85
		US-A-	5182204	26-01-93
		US-A-	5204015	20-04-93